

## CLAIMS

What is claimed is:

- 1 1. A method for synthesizing an auditory scene, comprising the steps of:  
2 (a) dividing an input audio signal into a plurality of different frequency bands; and  
3 (b) applying two or more different sets of one or more spatial parameters to two or more of the  
4 different frequency bands in the input audio signal to generate two or more synthesized audio signals of  
5 the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set  
6 of one or more spatial parameters is applied to the input audio signal as if the input audio signal  
7 corresponded to a single audio source in the auditory scene.
- 1 2. The invention of claim 1, wherein each set of one or more spatial parameters corresponds to a  
2 different audio source in the auditory scene.
- 1 3. The invention of claim 1, wherein, for at least one of the sets of one or more spatial parameters,  
2 at least one of the spatial parameters corresponds to a combination of two or more different audio sources  
3 in the auditory scene that takes into account relative dominance of the two or more different audio  
4 sources in the auditory scene.
- 1 4. The invention of claim 1, wherein the input audio signal is a mono signal.
- 1 5. The invention of claim 4, wherein the mono signal corresponds to a combination of two or more  
2 different mono source signals, wherein the two or more different frequency bands are selected by  
3 comparing magnitudes of the two or more different mono source signals, wherein, for each of the two or  
4 more different frequency bands, one of the mono source signals dominates the other mono source signals.
- 1 6. The invention of claim 4, wherein the mono signal corresponds to a combination of left and right  
2 audio signals of a binaural signal, wherein each different set of one or more spatial parameters is  
3 generated by comparing the left and right audio signals in a corresponding frequency band.
- 1 7. The invention of claim 1, wherein step (a) comprises the step of dividing the input audio signal  
2 into the plurality of different frequency bands based on information corresponding to the different sets of  
3 one or more spatial parameters.

1 8. The invention of claim 1, wherein each set of one or more spatial parameters is applied to at least  
2 one frequency band in which the input audio signal is dominated by a corresponding audio source in the  
3 auditory scene.

1 9. The invention of claim 1, wherein each set of one or more spatial parameters comprises one or  
2 more of an interaural level difference, an interaural time delay, and a head-related transfer function.

1 10. The invention of claim 1, wherein:

2 step (a) further comprises the step of converting the input audio signal from a time domain into a  
3 frequency domain; and

4 step (b) further comprises the step of converting the two or more synthesized audio signals from the  
5 frequency domain into the time domain.

1 11. The invention of claim 1, wherein the two or more synthesized audio signals comprise left and  
2 right audio signals of a binaural signal corresponding to the auditory scene.

1 12. The invention of claim 1, wherein the two or more synthesized audio signal comprise two or  
2 more signals of a multi-channel audio signal corresponding to the auditory scene.

1 13. The invention of claim 1, wherein:

2 the input audio signal is a mono signal;

3 each set of one or more spatial parameters corresponds to a different audio source in the auditory  
4 scene;

5 step (a) comprises the steps of:

6 (1) converting the mono signal from a time domain into a frequency domain;

7 (2) dividing the converted mono signal into the plurality of different frequency bands based on  
8 information corresponding to the sets of one or more spatial parameters;

9 each set of one or more spatial parameters is applied to at least one frequency band in which the input  
10 audio signal is dominated by a corresponding audio source in the auditory scene;

11 each set of one or more spatial parameters comprises one or more of an interaural level difference, an  
12 interaural time delay, and a head-related transfer function;

13 the two or more synthesized audio signals comprise left and right audio signals of a binaural signal  
14 corresponding to the auditory scene; and

15 step (b) further comprises the step of converting the left and right audio signals from the frequency  
16 domain into the time domain.

1 14. The invention of claim 13, wherein the mono signal corresponds to a combination of two or more  
2 different mono source signals, wherein the two or more different frequency bands are selected by  
3 comparing magnitudes of the two or more different mono source signals, wherein, for each of the two or  
4 more different frequency bands, one of the mono source signals dominates the other mono source signals.

1 15. The invention of claim 13, wherein the mono signal corresponds to a combination of left and  
2 right audio signals of a binaural signal, wherein each different set of one or more spatial parameters is  
3 generated by comparing the left and right audio signals in a corresponding frequency band.

1 16. A machine-readable medium, having encoded thereon program code, wherein, when the program  
2 code is executed by a machine, the machine implements a method for synthesizing an auditory scene,  
3 comprising the steps of:

4 (a) dividing an input audio signal into a plurality of different frequency bands; and

5 (b) applying two or more different sets of one or more spatial parameters to two or more of the  
6 different frequency bands in the input audio signal to generate two or more synthesized audio signals of  
7 the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set  
8 of one or more spatial parameters is applied to the input audio signal as if the input audio signal  
9 corresponded to a single audio source in the auditory scene.

1 17. An apparatus for synthesizing an auditory scene, comprising:

2 (a) means for dividing an input audio signal into a plurality of different frequency bands; and

3 (b) means for applying two or more different sets of one or more spatial parameters to two or more  
4 of the different frequency bands in the input audio signal to generate two or more synthesized audio  
5 signals of the auditory scene, wherein for each of the two or more different frequency bands, the  
6 corresponding set of one or more spatial parameters is applied to the input audio signal as if the input  
7 audio signal corresponded to a single audio source in the auditory scene.

1 18. An apparatus for synthesizing an auditory scene, comprising:

2 (1) an auditory scene synthesizer configured to:

3 (a) divide an input audio signal into a plurality of different frequency bands; and

4 (b) apply two or more different sets of one or more spatial parameters to two or more of the  
5 different frequency bands in the input audio signal to generate two or more synthesized audio signals of  
6 the auditory scene, wherein for each of the two or more different frequency bands, the corresponding set  
7 of one or more spatial parameters is applied to the input audio signal as if the input audio signal  
8 corresponded to a single audio source in the auditory scene; and

9 (2) one or more inverse time-frequency transformers configured to convert the two or more  
10 synthesized audio signals from a frequency domain into a time domain.

1 19. A method for processing two or more input audio signals, comprising the steps of:

2 (a) converting the two or more input audio signals from a time domain into a frequency domain;

3 (b) generating a set of one or more auditory scene parameters for each of two or more different  
4 frequency bands in the two or more converted input audio signals, where each set of one or more  
5 auditory scene parameters is generated as if the corresponding frequency band corresponded to a single  
6 audio source in an auditory scene; and

7 (c) combining the two or more input audio signals to generate a combined audio signal.

1 20. The invention of claim 19, wherein:

2 the two or more input audio signals are mono signals corresponding to different audio sources in the  
3 auditory scene;

4 each set of one or more auditory scene parameters corresponds to an audio source that dominates the  
5 other audio sources in the corresponding frequency band; and

6 the two or more input audio signals are combined in the time domain to generate the combined audio  
7 signal.

1 21. The invention of claim 19, wherein:

2 the two or more input audio signals are left and right audio signals of a binaural signal;

3 each set of one or more auditory scene parameters is generated by comparing the left and right audio  
4 signals in the corresponding frequency band;

5 the combined audio signal is generated by performing auditory scene removal on the left and right  
6 audio signals in the frequency domain based on the two or more sets of one or more auditory scene  
7 parameters; and

8 further comprising the step of converting the combined audio signal from the frequency domain into  
9 the time domain.

1 22. A machine-readable medium, having encoded thereon program code, wherein, when the program  
2 code is executed by a machine, the machine implements a method for processing two or more input audio  
3 signals, comprising the steps of:

- 4 (a) converting the two or more input audio signals from a time domain into a frequency domain;
- 5 (b) generating a set of one or more auditory scene parameters for each of two or more different  
6 frequency bands in the two or more converted input audio signals, where each set of one or more  
7 auditory scene parameters is generated as if the corresponding frequency band corresponded to a single  
8 audio source in an auditory scene; and
- 9 (c) combining the two or more input audio signals to generate a combined audio signal.

1 23. An apparatus for processing two or more input audio signals, comprising:

- 2 (a) means for converting the two or more input audio signals from a time domain into a frequency  
3 domain;
- 4 (b) means for generating a set of one or more auditory scene parameters for each of two or more  
5 different frequency bands in the two or more converted input audio signals, where each set of one or  
6 more auditory scene parameters is generated as if the corresponding frequency band corresponded to a  
7 single audio source in an auditory scene; and
- 8 (c) means for combining the two or more input audio signals to generate a combined audio signal.

1 24. An apparatus for processing two or more input audio signals, comprising:

- 2 (a) a time-frequency transformer configured to convert the two or more input audio signals from a  
3 time domain into a frequency domain;
- 4 (b) an auditory scene parameter generator configured to generate a set of one or more auditory scene  
5 parameters for each of two or more different frequency bands in the two or more converted input audio  
6 signals, where each set of one or more auditory scene parameters is generated as if the corresponding  
7 frequency band corresponded to a single audio source; and
- 8 (c) a combiner configured to combine the two or more input audio signals to generate a combined  
9 audio signal.

1 25. The invention of claim 24, wherein:

2 the two or more input audio signals are mono signals corresponding to different audio sources in the  
3 auditory scene;

4 each set of one or more auditory scene parameters corresponds to an audio source that dominates the  
5 other audio sources in the corresponding frequency band; and

